

A Distinguished Vacuum State for a Quantum Field in a Curved Spacetime: Formalism, Features, and Cosmology

Niayesh Afshordi,^{a,b} Siavash Aslanbeigi^{a,b} and Rafael D. Sorkin^{a,c}

^a*Perimeter Institute for Theoretical Physics, 31 Caroline St. N., Waterloo, ON, N2L 2Y5, Canada*

^b*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, N2L 3G1, Canada*

^c*Department of Physics, Syracuse University, Syracuse, NY 13244-1130, U.S.A.*

E-mail: nafshordi@perimeterinstitute.ca,

saslanbeigi@perimeterinstitute.ca, rsorkin@perimeterinstitute.ca

ABSTRACT: We define a distinguished “ground state” or “vacuum” for a free scalar quantum field in a globally hyperbolic region of an arbitrarily curved spacetime. Our prescription is motivated by the recent construction [1, 2] of a quantum field theory on a background causal set using only knowledge of the retarded Green’s function. We generalize that construction to continuum spacetimes and find that it yields a distinguished *vacuum* or *ground state* for a non-interacting, massive or massless scalar field. This state is defined for all compact regions and for many noncompact ones. In a static spacetime we find that our vacuum coincides with the usual ground state. We determine it also for a radiation-filled, spatially homogeneous and isotropic cosmos, and show that the super-horizon correlations are approximately the same as those of a thermal state. Finally, we illustrate the inherent non-locality of our prescription with the example of a spacetime which sandwiches a region with curvature in-between flat initial and final regions.

ARXIV EPRINT: [1205.1296](https://arxiv.org/abs/1205.1296)

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 2 | Background | 3 |
| 3 | The S-J Vacuum | 5 |
| 4 | Consistency with Known Vacua: Static Spacetimes | 10 |
| 5 | Application to Non-stationary Spacetimes | 12 |
| 5.1 | Massless field in the radiation era | 13 |
| 5.2 | Massive field in the radiation era | 14 |
| 6 | Causality and the S-J Vacuum | 17 |
| 7 | Conclusions and Discussions | 20 |
| A | Comments and Calculations | 21 |
| A.1 | When is $i\Delta$ self-adjoint? | 21 |
| A.2 | The S-J Vacuum and the Simple Harmonic Oscillator | 22 |
| A.3 | Equation (3.4) as an equality between distributions | 23 |
| A.4 | Another Slant on the SJ Prescription | 24 |
| B | Useful Theorems | 25 |

1 Introduction

The framework known as “quantum field theory in curved spacetime” concerns the interaction of quantum fields with gravity, but only in an asymmetrical sense. Non-gravitational, “matter” fields are treated in accord with quantum principles while their gravitational “back reaction” is either ignored entirely or described by a semiclassical form of the Einstein equations. Although not a fundamental theory of nature, this framework has provided us with profound insights into an eventual theory of quantum gravity. Important examples include Hawking radiation by black holes [3], the Unruh effect [4, 5], and the generation of Gaussian-distributed random perturbations in the theory of cosmic inflation [6]. In all these examples a choice of vacuum — or at least a reasonable reference state of the field — is of crucial importance. It therefore seems unsatisfactory that as it stands, quantum field theory lacks a general notion of “vacuum” which extends very far beyond flat spacetime.

Formulations of quantum field theory in Minkowski spacetime do provide a distinguished vacuum, but it rests heavily on a particle interpretation of the field that is closely

tied to the properties of the Fourier transform and the availability of plane waves. More abstract treatments tend to trace the uniqueness of the vacuum to Poincaré-invariance, but that is tied even more closely to flat space. It is thus unclear how one might extend the notion of vacuum beyond the case of spacetimes with a high degree of symmetry. (Moreover, even a large symmetry-group does not always yield a unique vacuum without further input. In de-Sitter for example, one has the one complex-parameter family of “ α -vacua” [7], all of which are invariant under the full de-Sitter group. To single one value of α out from the rest, one needs to impose the further condition that the two-point function take the so-called Hadamard form.)

One might even question whether a quantum field theory is well-defined at all before a vacuum is specified. What is probably the best studied mathematical framework for quantum field theory in flat space, that of the Wightman axioms, incorporates assertions about the vacuum among its basic assumptions, and it relies on them in proving such central results as the PCT and spin-statistics theorems. It is therefore noteworthy that the so-called algebraic approach to quantum field theory has been able to proceed a great distance without relying on a notion of vacuum, or indeed any unique representation of the quantum fields at all. In place of a Poincaré-invariant vacuum, it has been proposed to rely on a distinguished *class* of states, the so-called Hadamard states (which are well-suited to renormalization of the stress-tensor by “point-splitting”), supplemented by an assumption about a short-distance asymptotic expansion for products of quantum fields, namely the operator product expansion or “OPE” (see for example [8], [9] and references therein). If such a “purely algebraic” approach were to establish itself more generally, it might diminish the interest in distinguished “vacua” for curved spacetimes. Conversely, if a reasonable definition of a preferred vacuum state could be obtained it might remove some of the motivation for a purely algebraic formulation of quantum field theory.¹

Let us remark also that histories-based formulations of quantum mechanics tend to fuse the concept of state with that of equation of motion. This shows up clearly in formulations that start from the “quantum measure” [10–12] or “decoherence functional” [13], neither of which can be defined without furnishing a suitable set of “initial conditions”. In this sense, one has no dynamical law at all before a distinguished “initial state” is specified.

At a less formal level, the ability to think in terms of particles offers an obvious benefit to one’s intuition. And, especially in relation to cosmology, great interest attaches to the question whether certain sorts of states can be regarded as “natural” to certain regions of spacetime, a question we return to briefly in section 5.2. These, then, are two more reasons why the availability of a distinguished vacuum could be welcome, whether or not it is logically necessary to quantum field theory as such.

Moreover, what is logically necessary can change drastically if one passes from the spacetime continuum to some more fundamental structure, especially if that structure is discrete. As we will review later, the entire quantization process — as usually conceived — boils down to selecting an appropriate subspace of the solution space of the Klein-Gordon

¹We suspect that lasting enlightenment about the “best” formulation of quantum field theory will only arrive together with a solution of the problem of quantum gravity, by means of a greater theory within which that of quantum field theory in curved spacetime will have to be subsumed.

equation. But that way of organizing the problem seems to break down in the case of a causal set. There, the notion of “approximate solution” seems to be the best that is available, and one therefore requires a different starting point.

In [1], such a starting point was found in (the discrete analog of) the retarded Green function. On that basis a complete counterpart of the quantum field theory of a free scalar field was built up, and a unique “vacuum state” was derived. Herein, we generalize that derivation to quantum fields on continuum spacetimes, showing thereby that there is a sensible way to uniquely define a vacuum state for a scalar field in any globally hyperbolic spacetime or region of spacetime. More precisely, we consider the case of a *free scalar field* in a *globally hyperbolic* spacetime or region of spacetime, and in that context we put forward a definition of *distinguished vacuum state* that applies to all compact regions and to a large class of noncompact regions.

It is thus possible to carry the concept of vacuum far beyond the confines of Minkowski space by means of definitions we expose in detail below. Although, for all of the reasons indicated above, this possibility is of interest in itself, one naturally wants to know to what extent, and in what sense, our proposal is “the right one”? To that question, only a sufficient number of particular instances of our vacuum would seem to be germane. The examples of Minkowski spacetime and of globally static spacetimes furnish important evidence, but they contain little that is new physically. To judge the ultimate fruitfulness of our prescription, one should, for example, test it against the behavior of the “matter fields” that one actually encounters in the early universe. In Section 5 we make a start on this kind of test, beginning with the case of a spatially homogeneous and isotropic cosmology.

2 Background

In this section, we briefly review the quantization, along traditional lines, of a free scalar field on a curved spacetime. We will be a bit careful with the mathematical technicalities because it will benefit us later.² Consider a free, real-valued scalar field ϕ on a globally hyperbolic spacetime $(M, g_{\mu\nu})$ satisfying the Klein-Gordon equation with mass-parameter $m \geq 0$:

$$\nabla^\mu \nabla_\mu \phi(x) + m^2 \phi(x) = 0, \quad (2.1)$$

where ∇_μ is the covariant derivative operator on M .³ The condition of global hyperbolicity ensures that (2.1) has a well posed initial-value formulation (see theorem B.1 in Appendix B).

Let us review some mathematical structures that are important for both the classical and quantal theories of a free field. Consider a foliation of M by spacelike Cauchy surfaces Σ_t , labeled by a time parameter t . Let \mathbf{S} be the space of all real C^∞ solutions of (2.1) which induce initial data of compact support on some (and therefore on every) Σ_t . (This restriction on the solutions is just for convenience, so that various mathematical structures are well-defined.) The retarded and advanced Green’s functions $G_{R,A}(x, y)$ associated with

²Much of the discussion here will follow that of [8] and [14].

³We use signature $(+ - - -)$ and set $\hbar = c = 1$.

(2.1) satisfy

$$(\nabla^\mu \nabla_\mu + m^2)G_{R,A}(x, y) = -\frac{\delta^4(x - y)}{\sqrt{-g}}, \quad (2.2)$$

where g is the determinant of the metric-tensor. By definition $G_R(x, y) = 0$ unless $x \succ y$ (meaning x is inside or on the future lightcone of y), and $G_A(x, y) = 0$ unless $x \prec y$.

The so-called Pauli-Jordan function is defined as

$$\Delta(x, y) \equiv G_R(x, y) - G_A(x, y) = G_R(x, y) - G_R(y, x).^4 \quad (2.3)$$

From it we define an integral operator Δ :

$$(\Delta f)(x) \equiv \int_M \Delta(x, y) f(y) dV_y, \quad (2.4)$$

where $dV_y = \sqrt{-g(y)} d^4 y$ is the metric volume element on M , and we take the domain of Δ to be the space $C_0^\infty(M)$ of all smooth functions of compact support on M . Since Δf is the difference between two solutions of the inhomogeneous Klein-Gordon equation with the same source f , it satisfies the homogeneous Klein-Gordon equation (2.1):

$$(\square + m^2)\Delta = 0, \quad (2.5)$$

where $\square = \nabla_\mu \nabla^\mu$. Moreover, since f has compact support, Δf induces smooth initial data of compact support on all Cauchy surfaces, making Δ a map from $C_0^\infty(M)$ to \mathbf{S} . The operator Δ (or more generally the corresponding quadratic form) will be of crucial importance to us. A symplectic structure $\Omega : \mathbf{S} \times \mathbf{S} \rightarrow \mathbb{R}$ can be defined on \mathbf{S} :

$$\Omega(\phi_1, \phi_2) \equiv \int_{\Sigma_t} [\phi_1 \nabla_\mu \phi_2 - \phi_2 \nabla_\mu \phi_1] dS^\mu, \quad (2.6)$$

where $dS_\mu = n_\mu \sqrt{-h} d^3 x$ with n^a the unit normal to Σ_t , and h the determinant of the induced metric on Σ_t . The righthand side of (2.6) is well defined because it is independent of t for all solutions in \mathbf{S} .

To pass to the quantum theory, one introduces operator-valued distributions $\widehat{\phi}(x)$ that satisfy the Klein-Gordon equation, and the canonical commutation relations (CCR)

$$[\widehat{\phi}(f), \widehat{\phi}(g)] = i\Omega(\Delta f, \Delta g) = i \int_M f(x) \Delta(x, y) g(y) dV_x dV_y, \quad (2.7)$$

for all $f, g \in C_0^\infty(M)$, where $\widehat{\phi}(f) = \int_M \widehat{\phi}(x) f(x) dV_x$. The second equality follows from theorem B.2 (see Appendix B). Equation (2.7) is typically expressed as

$$[\widehat{\phi}(x), \widehat{\phi}(y)] = i\Delta(x, y). \quad (2.8)$$

To obtain operators in Hilbert space, one requires further a $*$ -representation of these relations. One typically works with irreducible, Fock representations constructed as follows:

- Complexify the Klein-Gordon solution space to get $\mathbf{S}_\mathbb{C} = \{\phi_1 + i\phi_2 | \phi_1, \phi_2 \in \mathbf{S}\}$.

⁴See Theorem B.3 of Appendix B.

- Define a map $(\cdot, \cdot)_{KG} : \mathbf{S}_{\mathbb{C}} \times \mathbf{S}_{\mathbb{C}} \rightarrow \mathbb{C}$ by $(\phi_1, \phi_2)_{KG} = i\Omega(\bar{\phi}_1, \phi_2)$, where the bar denotes complex conjugation. This map enjoys all the properties of a Hermitian inner product except that it's not positive definite.
- Choose any subspace $\mathcal{H} \subset \mathbf{S}_{\mathbb{C}}$ with the following properties:
 - The inner product $(\cdot, \cdot)_{KG}$ is positive definite on \mathcal{H} , thus making \mathcal{H} into a (pre-)Hilbert space over \mathbb{C} .
 - $\mathbf{S}_{\mathbb{C}}$ is equal to the span of \mathcal{H} and its complex conjugate space $\bar{\mathcal{H}}$.
 - For all $\phi^+ \in \mathcal{H}$ and $\phi^- \in \bar{\mathcal{H}}$, we have $(\phi^+, \phi^-)_{KG} = 0$.⁵

The Hilbert space is then taken to be the symmetric Fock space associated with \mathcal{H} , the field operators being defined as

$$\hat{\phi}(x) = \sum_i \psi_i(x) \hat{a}_i + \bar{\psi}_i(x) \hat{a}_i^\dagger, \quad (2.9)$$

where $\{\psi_i(x)\}$ is any orthonormal basis of (the Cauchy-completed) \mathcal{H} with respect to the inner product $(\cdot, \cdot)_{KG}$, and where $\{\hat{a}_i\}$ are the annihilation operators associated with $\{\psi_i\}$, satisfying the usual commutation relations, $[\hat{a}_i, \hat{a}_j] = 0$, $[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}$. To show that (2.9) satisfies the CCR, write $\hat{\phi}(f) = \sum_i (\int_M \psi_i f) \hat{a}_i + (\int_M \bar{\psi}_i f) \hat{a}_i^\dagger = -i \sum_i (\Delta f, \psi_i)_{KG} \hat{a}_i + (\Delta f, \bar{\psi}_i)_{KG} \hat{a}_i^\dagger$, where in the last equality we have used Theorem B.2. Then

$$\begin{aligned} [\hat{\phi}(f), \hat{\phi}(g)] &= - \sum_i (\Delta f, \psi_i)_{KG} (\Delta g, \bar{\psi}_i)_{KG} - (\Delta f, \bar{\psi}_i)_{KG} (\Delta g, \psi_i)_{KG} \\ &= (\Delta f) \sum_i |\psi_i\rangle_{KG} \langle \psi_i| - |\bar{\psi}_i\rangle_{KG} \langle \bar{\psi}_i| (\Delta g) \\ &= (\Delta f, \Delta g)_{KG} = i\Omega(\Delta f, \Delta g), \end{aligned} \quad (2.10)$$

where we have used the fact that $\sum_i |\psi_i\rangle_{KG} \langle \psi_i| - |\bar{\psi}_i\rangle_{KG} \langle \bar{\psi}_i|$ is the identity operator on $\mathbf{S}_{\mathbb{C}}$ (because $\{\psi_i\}$'s satisfy $(\psi_i, \psi_j)_{KG} = \delta_{ij}$, $(\psi_i, \bar{\psi}_j)_{KG} = 0$, and $(\bar{\psi}_i, \bar{\psi}_j)_{KG} = -\delta_{ij}$). Finally, the vacuum is defined as the state annihilated by all \hat{a}_i : $\hat{a}_i |0\rangle = 0$.

The trouble, of course, is that the subspace $\mathcal{H} \subset \mathbf{S}_{\mathbb{C}}$ is not unique. Even if we limit ourselves to Fock representations, there are many ways to choose \mathcal{H} , and with each one comes a different set of operators \hat{a}_i and a different vacuum.

3 The S-J Vacuum

The kernel $i\Delta(x, y)$ defined by (2.3) has two basic properties:

- Antisymmetric — because $G_A(x, y) = G_R(y, x)$,⁶
- Hermitian — i.e. $\overline{i\Delta(y, x)} = i\Delta(x, y)$.

⁵From here on, when we refer to “a basis $\{\psi_i\}$ of the Klein-Gordon solution space”, we mean that $\{\psi_i\}$ is an orthonormal basis of $\mathcal{H} \subset \mathbf{S}_{\mathbb{C}}$ with the above properties.

⁶See Theorem B.3 of Appendix B.

Let $L^2(M)$ denote the Hilbert space of all square-integrable functions on M ⁷ with the usual inner product

$$\langle \phi, \psi \rangle = \int_M \bar{\phi}(x) \psi(x) dV_x. \quad (3.1)$$

Then, the integral operator $i\Delta$ associated with $i\Delta(x, y)$ [defined in (2.4)] will be Hermitian on a subspace of $L^2(M)$: $\langle \phi, i\Delta\psi \rangle = \langle i\Delta\phi, \psi \rangle$ for all ϕ and ψ in the domain of $i\Delta$.⁸ Strictly speaking, $i\Delta$ is in general only a densely defined quadratic form in $L^2(M)$, but in introducing our prescription of a vacuum, let us at first set aside the functional-analytical subtleties associated with the domain of $i\Delta$ and assume it to be a self-adjoint operator on $L^2(M)$, so that $i\Delta(x, y)$ can be “diagonalized” in the sense of the spectral theorem. We can now state our prescription for a distinguished state, which we call the *S-J vacuum* after the authors of [1, 2].

Assuming that $i\Delta$ is selfadjoint, a vacuum state $|SJ\rangle$ can be defined covariantly via $\langle SJ | \hat{\phi}(x) \hat{\phi}(y) | SJ \rangle = \text{Pos}(i\Delta(x, y))$, where $\text{Pos}(i\Delta(x, y))$ is the positive spectral projection of $i\Delta$ provided by the spectral theorem.

Informally speaking, this just means the following: (i) construct $\Delta(x, y)$ by anti-symmetrizing the retarded Green’s function, which is uniquely determined from the Klein-Gordon equation in any globally hyperbolic spacetime (see theorem B.1 in Appendix B), (ii) diagonalize $i\Delta(x, y)$ in the L^2 norm, (iii) take its ‘positive part’ to be the two-point function. The spectral theorem gives precise mathematical sense to this prescription so long as $i\Delta$ is self-adjoint.

Of course a two-point function is not yet a full characterization of a state, but it becomes so if we take the state to be “gaussian” by appropriately expressing the n -point functions in terms of the two-point function (i.e. by means of the Wick rule). It then will follow, given (2.9) and (3.2) below, that the Gel’fand-Naimark-Segal (GNS) representation associated with $|SJ\rangle$ will be a Fock representations with $|SJ\rangle$ playing the role of vacuum.

Let us consider first the special case where $i\Delta(x, y) \in L^2(M \times M)$,⁹ as is the case in a bounded¹⁰ globally hyperbolic region of 1+1 dimensional Minkowski space, for instance.¹¹ Then $i\Delta$ becomes a so-called Hilbert-Schmidt integral operator and the following version of the spectral theorem applies [15]: there exists an orthonormal basis $\{T_k\}_{k=1}^\infty$ of $L^2(M)$ consisting of eigenfunctions of $i\Delta$ which satisfy $i\Delta T_k = \lambda_k T_k$ with $\lambda_k \in \mathbb{R}$. Using this theorem and the fact that $\Delta(x, y)$ itself is real, we deduce $i\Delta \bar{T}_k = -\lambda_k \bar{T}_k$, which in turn

⁷ $L^2(M) = \{\psi : M \rightarrow \mathbb{C} \mid \int_M |\psi(x)|^2 dV_x < \infty\}$.

⁸ $\langle \phi, i\Delta\psi \rangle = \int_M \int_M i\Delta(x, y) \bar{\phi}(x) \psi(y) dV_x dV_y = - \int_M \int_M i\Delta(y, x) \bar{\phi}(x) \psi(y) dV_x dV_y = \int_M \int_M i\Delta\phi(y) \psi(y) dV_y = \langle i\Delta\phi, \psi \rangle$.

⁹ i.e. $\int_M \int_M |i\Delta(x, y)|^2 dV_x dV_y < \infty$.

¹⁰ Bounded = having compact closure.

¹¹ For a massive scalar field in 1+1 dimensional Minkowski space with metric $ds^2 = dt^2 - dx^2$, $G_R(t, x; t', x') = G(t - t'; x - x')$ where $G(t, x) = \theta(t) \theta(s^2)^{\frac{1}{2}} J_0(ms)$, J_0 is a Bessel function of the first kind, and $s^2 = t^2 - x^2$ [1].

makes it possible to split $i\Delta(x, y)$ into a *positive* and a *negative* part:

$$i\Delta(x, y) = \sum_{k=1}^{\infty} \lambda_k \{T_k(x)\bar{T}_k(y) - \bar{T}_k(x)T_k(y)\}$$

(taking $\lambda_k > 0$ now). In this case, our prescription can be expressed as

$$W(x, y) \equiv \langle SJ | \hat{\phi}(x)\hat{\phi}(y) | SJ \rangle = \sum_{k=1}^{\infty} \lambda_k T_k(x)\bar{T}_k(y). \quad (3.2)$$

This is equivalent to introducing field operators $\hat{\phi}(x) = \sum_k \sqrt{\lambda_k} \{T_k(x)\hat{a}_k + \bar{T}_k(x)\hat{a}_k^\dagger\}$, because (i) eigenfunctions of $i\Delta$ with $\lambda \neq 0$ necessarily satisfy the Klein-Gordon equation¹² and (ii) the commutation relations are trivially satisfied.¹³ The S-J vacuum is then the state in Fock space that is annihilated by all \hat{a}_k .

Now let us turn to 3+1 dimensions, where $i\Delta$ is no longer Hilbert-Schmidt. Nonetheless, $i\Delta(x, y)$ is still a distribution and, at least within Minkowski space, it defines a self-adjoint operator if M is bounded (see section A.1 of Appendix A). Thanks to the spectral theorem, our prescription then retains a precise mathematical sense, and it is not too far-fetched to assume that this continues to hold for curved spacetimes, because curvature should not change the singularity structure of $i\Delta(x, y)$ too drastically.¹⁴

Although selfadjointness might seem to be merely a technical issue, it highlights the fact that the S-J vacuum depends on a choice of (globally hyperbolic) spacetime region. Indeed, as we have just seen, our prescription is not guaranteed to be well defined unless one chooses a region that is bounded, both spatially and temporally. Thus arise two questions: To what extent does our prescription depend on boundary conditions, and to what extent does it remain well-defined in unbounded spacetimes?

In answering the first question, one must distinguish between spatial boundaries (also referred to as timelike boundaries) and temporal ones. Spatial boundaries are familiar to us from putting fields in a box, Casimir effect, etc; and they seem unproblematic. When they are present the S-J vacuum will be sensitive to one's choice of boundary conditions, because the retarded and advanced Green functions do depend on them. But this is as it should be since the physics genuinely depends on the boundary conditions. We will also consider below regions which are unbounded spatially, but no special difficulties will arise from that feature.

The case of a temporal boundary (spacelike or null) is less familiar. The first thing to notice is that *boundary conditions are neither needed nor possible* in this case, since the region is (by assumption) globally hyperbolic. Mathematically, this very satisfactory feature stems from the fact that Δ is not a differential operator but an integral one. Nevertheless, one must bear in mind that the “ground state” one ends up with, does

¹² $i\Delta f = \lambda f \implies (\square + m^2)f = (i/\lambda)(\square + m^2)\Delta f = 0$ since $(\square + m^2)\Delta = 0$.

¹³Note that in this case there is no need to ‘smear out’ the field operators with smooth test-functions of compact support, because $i\Delta(x, y)$ is a completely well-defined function (in the L^2 sense).

¹⁴Fewster and Verch have now established rigorously that our proposal is well-defined on all bounded globally hyperbolic spacetimes and that the S-J vacuum is a “pure quasi-free state” [16].

depend on the region with which one begins. In itself, this dependence on the region merely expresses the nonlocal character of our definition. One might for example be interested in which “vacuum” would be appropriate to an early stage of expansion of the cosmos, and one would not want in that case to apply our definition to the full spacetime, including its whole future development. However, one might also want to apply the definition to unbounded spacetimes like Minkowski space, and in such cases one needs to worry about dependence on an eventual infrared cutoff. If the metric is static, for instance, why should time play any role in what the vacuum state looks like?

In dealing with such instances, it is always possible to work first in a truncated spacetime, and later send the temporal boundary to infinity. In section A.2 of Appendix A, we apply this method to the simple harmonic oscillator and show that it succeeds in the sense that the resulting S-J vacuum is the minimum energy state of the Hamiltonian. As we will later demonstrate, this continues to be true for all static spacetimes. Another example of such a calculation can be found in [17], where the spectrum of $i\Delta$ is computed in a 1+1 dimensional causal-diamond, and it is found that (up to the usual infrared ambiguities that affect massless scalars in 2d) the resulting two-point function has the correct (i.e. Minkowski) limiting behaviour as the boundaries of the diamond tend to infinity. However, there are also cases where taking a temporal cutoff to infinity is an ill-defined procedure. In [18], it is shown that the so-called Poincaré patch of de Sitter space provides an example of such a case. (We suspect that this kind of ambiguity can be understood intuitively as the failure of $i\Delta$ to admit a selfadjoint extension which is unique. However we don’t know how to pose such a question properly, because Δ is densely defined only as a quadratic form on $L^2(M)$, not as an operator from $L^2(M)$ to itself.)

Diagonalizing $i\Delta$ lies at the heart of our prescription, at least in practice. In this section, we will attempt to frame this problem as generally as possible. In subsequent sections, we will deal with more concrete examples. Let $\{u_k(x)\}$ be a basis for the Klein-Gordon solution space.¹⁵ Expanding the field operator in terms of these modes as $\hat{\phi}(x) = \sum_k u_k(x)\hat{a}_k + \bar{u}_k(x)\hat{a}_k^\dagger$ and computing the commutator yields, as we have seen,

$$[\hat{\phi}(x), \hat{\phi}(y)] = \sum_k u_k(x)\bar{u}_k(y) - u_k(y)\bar{u}_k(x). \quad (3.3)$$

It then follows from the CCR (2.8) that the integral-kernel $i\Delta(x, x')$ takes the form,¹⁶

$$i\Delta(x, y) = \sum_k u_k(x)\bar{u}_k(y) - \bar{u}_k(x)u_k(y). \quad (3.4)$$

Of course the choice of the $u_k(x)$ in this expansion is not unique. Another set of modefunctions $g_k(x) = \sum_n \alpha_{kn}u_n(x) + \beta_{kn}\bar{u}_n(x)$ will give the same commutator, so long as the following normalization conditions are met: $\sum_k \alpha_{ik}\bar{\alpha}_{jk} - \beta_{ik}\bar{\beta}_{jk} = \delta_{ij}$ and $\sum_k \alpha_{ik}\beta_{jk} - \beta_{ik}\alpha_{jk} = 0$, this being nothing but a Bogoliubov transformation.

¹⁵See footnote 5 on the definition of “basis”.

¹⁶In Section A.3 of Appendix A we confirm that the two sides of this equation are equal when integrated against an arbitrary test function.

By means of such a transformation, we can find orthonormal eigenfunctions $T_n(x)$ of $i\Delta$ (with corresponding eigenvalues λ_n) starting from any convenient basis $\{u_k(x)\}$ of the Klein-Gordon solution space. Requiring $T_n(x)$'s to be eigenfunctions of $i\Delta$ and using (3.4):

$$i\Delta T_n(x) = \lambda_n T_n(x) = \sum_k \langle u_k, T_n \rangle u_k(x) - \langle \bar{u}_k, T_n \rangle \bar{u}_k(x). \quad (3.5)$$

For notational simplicity, let $\alpha_{nk} \equiv \frac{\langle u_k, T_n \rangle}{\lambda_n}$ and $\beta_{nk} \equiv -\frac{\langle \bar{u}_k, T_n \rangle}{\lambda_n}$ so that $T_n(x) = \sum_k \alpha_{nk} u_k(x) + \beta_{nk} \bar{u}_k(x)$. These coefficients then satisfy:

$$\alpha_{nk} = \frac{1}{\lambda_n} \sum_m \alpha_{nm} \langle u_k, u_m \rangle + \beta_{nm} \langle u_k, \bar{u}_m \rangle \quad (3.6)$$

$$\beta_{nk} = \frac{-1}{\lambda_n} \sum_m \alpha_{nm} \langle \bar{u}_k, u_m \rangle + \beta_{nm} \langle \bar{u}_k, \bar{u}_m \rangle. \quad (3.7)$$

Requiring these eigenfunctions to be orthonormal ($\forall n, m$) yields

$$\langle T_n, \bar{T}_m \rangle = 0 \iff \sum_k \alpha_{nk} \beta_{mk} - \alpha_{mk} \beta_{nk} = 0 \quad (3.8)$$

$$\langle T_n, T_m \rangle = \delta_{nm} \iff \sum_k \bar{\alpha}_{nk} \alpha_{mk} - \bar{\beta}_{nk} \beta_{mk} = \frac{\delta_{nm}}{\lambda_m}. \quad (3.9)$$

Then, diagonalizing $i\Delta$ boils down to finding α_{nk} and β_{nk} by solving these four equations. (We have not addressed the issue of convergence in the sums appearing above. In fact, if it turned out that the $\{T_n(x)\}$ and $\{u_n(x)\}$ induce unitarily inequivalent representations of CCR, the above sums would not converge.)

It is not obvious how this can be done generically. To proceed, let us simplify this calculation by *assuming* that there are modefunctions $\{u_k(x)\}$ that satisfy

$$\langle u_k, u_m \rangle = \langle u_k, u_k \rangle \delta_{km} \quad (3.10)$$

$$\langle u_k, \bar{u}_m \rangle = \langle u_k, \bar{u}_{-k} \rangle \delta_{k, -m}. \quad (3.11)$$

The notation used here is as follows: for every $u_n \in \{u_k\}$, there is one (and only one) member of the complex conjugate set $\bar{u}_{-n} \in \{\bar{u}_k\}$, for which $\langle u_n, \bar{u}_{-n} \rangle$ can be non-zero. Also, we denote the complex conjugate of \bar{u}_{-n} by u_{-n} .

Our assumption is motivated by spacetimes for which these modefunctions are plane waves. Under this assumption, α_{nk} and β_{nk} can be found:

$$\alpha_{nk} = \left[\frac{\lambda_{-n} + \langle u_n, u_n \rangle}{\lambda_n(\lambda_n + \lambda_{-n})} \right]^{\frac{1}{2}} \delta_{nk} \quad (3.12)$$

$$\beta_{nk} = - \left[\frac{\langle u_n, u_n \rangle - \lambda_n}{\lambda_n(\lambda_n + \lambda_{-n})} \right]^{\frac{1}{2}} e^{-i\text{Arg}(\langle u_n, \bar{u}_{-n} \rangle)} \delta_{n, -k}, \quad (3.13)$$

where λ_n is given by

$$2\lambda_n = \langle u_n, u_n \rangle - \langle u_{-n}, u_{-n} \rangle + \left[(\langle u_n, u_n \rangle + \langle u_{-n}, u_{-n} \rangle)^2 - 4|\langle u_n, \bar{u}_{-n} \rangle|^2 \right]^{\frac{1}{2}}. \quad (3.14)$$

By direct substitution, it can be verified that these indeed solve (3.6), (3.7), (3.8) and (3.9). Finally, the vacuum state modefunctions as picked by our prescription $\mathcal{U}_n^{SJ}(x) \equiv \sqrt{\lambda_n} T_n(x)$ take the form

$$\begin{aligned} \sqrt{2}\mathcal{U}_n^{SJ}(x) = & \left[1 + \frac{1}{\sqrt{1 - \frac{4|\langle u_n, \bar{u}_{-n} \rangle|^2}{(\langle u_n, u_n \rangle + \langle u_{-n}, u_{-n} \rangle)^2}}} \right]^{\frac{1}{2}} u_n(x) \\ & - \left[-1 + \frac{1}{\sqrt{1 - \frac{4|\langle u_n, \bar{u}_{-n} \rangle|^2}{(\langle u_n, u_n \rangle + \langle u_{-n}, u_{-n} \rangle)^2}}} \right]^{\frac{1}{2}} e^{-i\text{Arg}(\langle u_n, \bar{u}_{-n} \rangle)} \bar{u}_{-n}(x). \end{aligned} \quad (3.15)$$

Looking closely at (3.15), we see that our prescription picks out a particular basis of the Klein-Gordon solution space that satisfies (3.10), (3.11), and

$$\frac{|\langle u_n, \bar{u}_{-n} \rangle|}{\langle u_n, u_n \rangle + \langle u_{-n}, u_{-n} \rangle} = 0. \quad (3.16)$$

In a bounded region of spacetime where all inner products are finite, (3.16) implies that $\langle u_n, \bar{u}_{-n} \rangle = 0$. That this is a unique choice can be shown on more general grounds (see section A.4 of Appendix A). Even in an unbounded spacetime, where the inner-products might diverge, we may deem (3.16) to be satisfied so long as the denominator diverges more strongly than the numerator. This is very similar to the case of the simple harmonic oscillator (see section A.2 of Appendix A). However, it might be the case that $\frac{2|\langle u_n, \bar{u}_{-n} \rangle|}{\langle u_n, u_n \rangle + \langle u_{-n}, u_{-n} \rangle} \rightarrow 1$ as the limit is taken to infinity. In such a case, the prescription (3.15) fails.

4 Consistency with Known Vacua: Static Spacetimes

In static spacetimes, i.e. spacetimes that admit an everywhere time-like and hypersurface-orthogonal Killing vector k^μ , a natural choice of vacuum modefunctions exists, namely the solutions of the Klein-Gordon equation which are purely positive frequency with respect to the Killing time. (The corresponding vacuum state minimizes the Hamiltonian). For a massive scalar field $\psi(x)$ in a static spacetime the Klein-Gordon equation reads

$$\left(\frac{\partial^2}{\partial t^2} + K \right) \psi(x) = 0, \quad (4.1)$$

where $K = \frac{1}{g^{00}(\vec{x})} \left\{ \frac{1}{\sqrt{-g(\vec{x})}} \partial_i [g^{ij}(\vec{x}) \sqrt{-g(\vec{x})} \partial_j] + m^2 \right\}$ is a purely spatial differential operator. Let $L^2(\Sigma)$ denote the Hilbert space of all L^2 functions on the spatial domain Σ with inner product

$$\langle f, g \rangle_S = \int_{\Sigma} \bar{f}(\vec{x}) g(\vec{x}) \frac{\sqrt{-h}}{|k|} d^3\vec{x}, \quad (4.2)$$

where $|k| = \sqrt{k_\mu k^\mu} = \sqrt{g_{00}}$, and h is the determinant of the induced metric on Σ . Let us assume that K is a self-adjoint and strictly positive operator on $L^2(\Sigma)$ so that it has a

well-defined positive spectrum: $(Ku_k)(\vec{x}) = \omega(k)^2 u_k(\vec{x})$. Then, it is always possible to find complex solutions of the Klein-Gordon equation of the form $\psi_k(x) = \frac{e^{-i\omega(k)t}}{\sqrt{2\omega(k)}} u_k(\vec{x})$ ($\omega > 0$) that satisfy

$$L_k \psi_k = k^\mu \nabla_\mu \psi_k = -i\omega \psi_k, \quad (4.3)$$

where $L_k \psi_n$ denotes the Lie derivative of ψ_n along the Killing vector k^μ . Moreover, in static spacetimes, we can choose $n^\mu = \frac{k^\mu}{|k|}$ as the unit normal to Σ . It then follows that $\{\psi_k(x)\}$ form an orthonormal basis of the Klein-Gordon solution space so long as $\langle u_k, u_{k'} \rangle_S = \delta_{kk'}$:

$$(\psi_k, \psi_{k'})_{KG} = \frac{\omega(k) + \omega(k')}{\sqrt{2\omega(k)}\sqrt{2\omega(k')}} e^{it(\omega(k) - \omega(k'))} \langle u_k, u_{k'} \rangle_S = \delta_{kk'} \quad (4.4)$$

$$(\bar{\psi}_k, \psi_{k'})_{KG} = \frac{\omega(k') - \omega(k)}{\sqrt{2\omega(k)}\sqrt{2\omega(k')}} e^{-it(\omega(k) + \omega(k'))} \langle \bar{u}_k, u_{k'} \rangle_S = 0. \quad (4.5)$$

¹⁷ Let us now turn to the S-J prescription. It can be verified that

$$\langle \psi_k, \psi_{k'} \rangle = \langle \sqrt{\frac{\pi g_{00}}{\omega(k)}} u_k, \sqrt{\frac{\pi g_{00}}{\omega(k')}} u_{k'} \rangle_S \delta(\omega(k) - \omega(k')) \quad (4.6)$$

$$\langle \psi_k, \bar{\psi}_{k'} \rangle = \langle \sqrt{\frac{\pi g_{00}}{\omega(k)}} u_k, \sqrt{\frac{\pi g_{00}}{\omega(k')}} \bar{u}_{k'} \rangle_S \delta(\omega(k) + \omega(k')) = 0, \quad (4.7)$$

where we have taken $t \in (-\infty, \infty)$. This implies that $i\Delta$ does not mix the positive and negative frequency modefunctions, which immediately follows from (3.4):

$$(i\Delta \psi_k)(x) = \sum_{k'} \langle \psi_{k'}, \psi_k \rangle \psi_{k'}(x). \quad (4.8)$$

This equation can be viewed as a matrix $\overline{\langle \psi_k, \psi_{k'} \rangle}$ multiplying a vector $\psi_{k'}(x)$, where $\overline{\langle \psi_k, \psi_{k'} \rangle}$ is *Hermitian* and can be diagonalized. As a result, it is always possible to find eigenfunctions of $i\Delta$ as linear combinations of *purely positive frequency modes*, which in turn implies that the S-J vacuum state is the same as the vacuum state defined by $\{\psi_k\}$'s. We have overlooked some potential technical difficulties in this argument. For instance, it might be the case that the inner products $\langle \sqrt{g_{00}} u_k, \sqrt{g_{00}} u_{k'} \rangle_S$ are infinite because the spatial domain is not compact. However, because in this case the eigenfunctions of $i\Delta$ are only linear combinations of purely positive frequency modes, the resulting two-point function will *not* depend on the regularization scheme, as it ought to be equal to $\langle SJ | \hat{\phi}(x) \hat{\phi}(y) | SJ \rangle = \sum_k \psi_k(x) \bar{\psi}_k(y)$. ¹⁸

Consider, for example, the 1+1 dimensional Rindler wedge:

$$ds^2 = e^{2a\xi} (d\eta^2 - d\xi^2), \quad (4.9)$$

where $-\infty < \eta, \xi < \infty$. With $m = 0$, the normalized positive frequency modes with respect to the Killing vector ∂_η are

$$\psi_k = \frac{1}{\sqrt{4\pi|k|}} e^{-i|k|\eta + ik\xi}. \quad (4.10)$$

¹⁷Note that $(\bar{\psi}_k, \psi_{k'})$ vanishes identically because $\langle \bar{u}_k, u_{k'} \rangle_S$ is only nonzero when \bar{u}_k and $u_{k'}$ both have the same eigenvalue, but in that case $\omega(k') - \omega(k) = 0$.

¹⁸The same argument applies to the $\delta(0)$'s that appear.

These modefunctions form an orthonormal basis of the Klein-Gordon solution space: $(\psi_k, \psi_{k'})_{KG} = \delta(k - k')$ and $(\psi_k, \bar{\psi}_{k'})_{KG} = 0$. Moreover,

$$\langle \psi_k, \psi_{k'} \rangle = \frac{\delta(|k| - |k'|)}{\sqrt{4|k||k'|}} \int_{-\Lambda}^{\Lambda} e^{i\xi(k' - k)} e^{2a\xi} d\xi \quad (4.11)$$

$$\langle \bar{\psi}_k, \psi_{k'} \rangle = \frac{\delta(|k| + |k'|)}{\sqrt{4|k||k'|}} \int_{-\Lambda}^{\Lambda} e^{i\xi(k' + k)} e^{2a\xi} d\xi = 0, \quad (4.12)$$

where we have regulated the spatial integrals with a cut-off Λ . In this case, all inner products vanish except for $\langle \psi_k, \psi_k \rangle$ and $\langle \psi_k, \psi_{-k} \rangle$ and it follows from (3.4) that $(i\Delta\psi_k)(x) = \langle \psi_k, \psi_k \rangle \psi_k(x) + \langle \psi_{-k}, \psi_k \rangle \psi_{-k}(x)$. Then, it can be verified that orthonormal eigenfunctions $T_k(x)$ of $i\Delta$ (with eigenvalue λ_k) take the form

$$T_k(x) = \frac{1}{\sqrt{2\lambda_k}} \left(\psi_k(x) + e^{i\text{Arg}\langle \psi_{-k}, \psi_k \rangle} \psi_{-k}(x) \right) \quad (4.13)$$

where

$$\lambda_k = \frac{1}{\langle \psi_k, \psi_k \rangle + |\langle \psi_{-k}, \psi_k \rangle|}. \quad (4.14)$$

Therefore, the S-J vacuum modefunctions are

$$\mathcal{U}_n^{SJ}(x) \equiv \sqrt{\lambda_n} T_n(x) = \frac{1}{\sqrt{2}} \left(\psi_k(x) + e^{i\text{Arg}\langle \psi_{-k}, \psi_k \rangle} \psi_{-k}(x) \right). \quad (4.15)$$

In this expression, $\text{Arg} \langle \psi_{-k}, \psi_k \rangle$ depends on the cut-off Λ used to regulate the spatial integrals. However, as previously argued, this makes no difference because the two-point function is independent of Λ .

It is worth noting that the foregoing analysis does *not* apply to stationary spacetimes that are not static, including cases where the Killing vector under consideration is not everywhere timelike. It would be particularly interesting to investigate the S-J vacuum in the spacetime of a rotating star with an ergo-region.

5 Application to Non-stationary Spacetimes

Quantum field theory on time-dependent backgrounds is of particular importance because the universe we live in is not static. The choice of vacuum in such cases is not at all trivial. For example, in a Friedmann-Lemaître-Robertson-Walker (FLRW) spacetime, one choice of instantaneous vacuum is obtained by minimizing the Hamiltonian at the given instant in time. This might seem like a natural generalization from static spacetimes, but as is well-known by now, it suffers from severe physical problems like infinite particle production [19]. In this section, we will work out the S-J vacuum state in a spatially-flat FLRW spacetime for some specific cases. (In [20], a prescription was introduced for singling out a Hadamard state in the case of a spatially-flat FLRW universe with a scale factor that is either exponential or a power-law at early times. We defer a comparison between the SJ state and that introduced in [20] to future work.) The metric reads

$$ds^2 = a(\eta)^2 [d\eta^2 - d\vec{x}^2], \quad (5.1)$$

where $a(\eta)$ and η are the scale factor and conformal time, respectively. A basis¹⁹ $\{u_k\}$ for the Klein-Gordon solution space may be constructed as

$$u_k(x) = \frac{e^{i\vec{k}\cdot\vec{x}}}{\sqrt{(2\pi)^3}} \frac{g_k(\eta)}{a(\eta)} \quad (5.2)$$

where k is the comoving Fourier wavenumber and $g_k(\eta)$ satisfies²⁰

$$\left(\partial_\eta^2 + k^2 + m^2 a^2 - \frac{a''}{a} \right) g_k(\eta) = 0 \quad (5.3)$$

$$g_k(\eta) \partial_\eta \bar{g}_k(\eta) - \bar{g}_k(\eta) \partial_\eta g_k(\eta) = i. \quad (5.4)$$

Satisfying (5.3) is equivalent to satisfying the Klein-Gordon equation, while the normalization of the Wronskian in (5.4) is equivalent to $(u_k, u_{k'})_{KG} = \delta^3(\vec{k} - \vec{k}')$. We have also $(u_k, \bar{u}_{k'})_{KG} = 0$. Moreover,

$$\langle u_k, u_{k'} \rangle = \langle g_k, g_{k'} \rangle_\eta \delta^3(\vec{k} - \vec{k}') \quad (5.5)$$

$$\langle u_k, \bar{u}_{k'} \rangle = \langle g_k, \bar{g}_{k'} \rangle_\eta \delta^3(\vec{k} + \vec{k}'), \quad (5.6)$$

where \langle, \rangle_η is defined by

$$\langle f(\eta), g(\eta) \rangle_\eta \equiv \int_0^\Lambda \bar{f}(\eta) g(\eta) a^2(\eta) d\eta. \quad (5.7)$$

As usual, we have regulated the integral with a cutoff Λ , which will be taken to infinity after the eigenfunctions of $i\Delta$ are found. We can now use (3.15) to compute the spectrum of $i\Delta$:

$$\mathcal{U}_k^{SJ}(x) = \frac{e^{i\vec{k}\cdot\vec{x}}}{\sqrt{2(2\pi)^3} a(\eta)} \left\{ \sqrt{1 + \frac{1}{\sqrt{1 - \frac{|\langle g_k, \bar{g}_k \rangle_\eta|^2}{\langle g_k, g_k \rangle_\eta^2}}}} g_k(\eta) - \sqrt{-1 + \frac{1}{\sqrt{1 - \frac{|\langle g_k, \bar{g}_k \rangle_\eta|^2}{\langle g_k, g_k \rangle_\eta^2}}}} e^{i \text{Arg} \langle \bar{g}_k, g_k \rangle} \bar{g}_k(\eta) \right\}. \quad (5.8)$$

5.1 Massless field in the radiation era

In the radiation era $a \propto \eta$. When $m = 0$, $g_k(\eta) = \frac{1}{\sqrt{2k}} e^{-ik\eta}$ satisfies both (5.3) and (5.4), and we have also

$$\frac{|\langle g_k, \bar{g}_k \rangle_\eta|}{\langle g_k, g_k \rangle_\eta} = \lim_{\Lambda \rightarrow \infty} \frac{|\int_0^\Lambda e^{2ik\eta} \eta^2 d\eta|}{\int_0^\Lambda \eta^2 d\eta} = 0. \quad (5.9)$$

Putting this back in (5.8), our prescription picks out the modefunctions:

$$\mathcal{U}_k^{SJ}(x) = \frac{1}{\sqrt{(2\pi)^3 2k} a(\eta)} e^{-i(k\eta - \vec{k}\cdot\vec{x})}. \quad (5.10)$$

These are the so-called adiabatic-vacuum modefunctions (for which an exact expression exists in the case of a massless scalar field in a radiation dominated cosmos) [21].

¹⁹See footnote 5 on the definition of “basis”.

²⁰ $k = |\vec{k}|$ and $a' = \frac{da(\eta)}{d\eta}$.

5.2 Massive field in the radiation era

For a massive free scalar field in the radiation era, (5.3) can still be solved analytically. Let $z = i\tilde{m}\eta^2$, where $\tilde{m} = \alpha m$ and α is a constant defined through $a(\eta) = \alpha\eta$. Furthermore, define a function F by $g_k(\eta) \equiv \frac{F(i\tilde{m}\eta^2)}{\sqrt{\eta}}$. With these definitions, (5.3) becomes

$$\frac{\partial^2 F}{\partial z^2} + \left(-\frac{1}{4} - \frac{ik^2}{4\tilde{m}z} + \frac{3}{16z^2}\right)F = 0. \quad (5.11)$$

This equation has two independent solutions $W_{\lambda,\mu}(z)$ and $W_{-\lambda,\mu}(-z)$ (called Whittaker functions) with $\lambda = \frac{-ik^2}{4\tilde{m}}$ and $\mu = \frac{1}{4}$ (see e.g. [22]). In our case, these two functions are complex conjugates of one another. Using the properties of Whittaker functions,²¹ it can be shown that g_k satisfies the Wronskian condition (5.4) with the normalization:

$$g_k(\eta) = \frac{e^{\frac{-\pi k^2}{8\tilde{m}}}}{\sqrt{2ma(\eta)^{1/2}}} W_{\lambda,\mu}(i\tilde{m}\eta^2). \quad (5.12)$$

As we will soon show, $|g_k(\eta)|$ is constant for small η , which means all inner products are finite in this region. Divergences arise for large η , though. In this regime, $W_{\lambda,\mu}(z) \rightarrow e^{-\frac{1}{2}z}z^\lambda$ and plugging this into (5.12) we find:

$$g_k(\eta) \rightarrow \frac{1}{\sqrt{2ma(\eta)^{1/2}}} e^{-\frac{i}{2}[\tilde{m}\eta^2 + \frac{k^2}{2\tilde{m}} \ln(\tilde{m}\eta^2)]}. \quad (5.13)$$

Just as before, it can be checked that $\lim_{\Lambda \rightarrow \infty} \frac{|\langle g_k, \bar{g}_k \rangle_\eta|}{\langle g_k, g_k \rangle_\eta} = 0$,²² whence our prescription picks out the modefunctions

$$\mathcal{U}_k^{SJ}(x) = \frac{e^{i\vec{k} \cdot \vec{x}} e^{\frac{-\pi k^2}{8\tilde{m}}}}{\sqrt{2m(2\pi)^3 a(\eta)^{3/2}}} W_{\frac{-ik^2}{4\tilde{m}}, \frac{1}{4}}(i\tilde{m}\eta^2). \quad (5.14)$$

The corresponding two-point function is:

$$\langle SJ | \hat{\phi}(x) \hat{\phi}(x') | SJ \rangle = \int \frac{d^3 \vec{k}}{(2\pi)^3} \frac{e^{i\vec{k} \cdot (\vec{x} - \vec{x}') - \frac{\pi k^2}{4\tilde{m}}}}{2ma(\eta)^{3/2} a(\eta')^{3/2}} W_{\frac{-ik^2}{4\tilde{m}}, \frac{1}{4}}(i\tilde{m}\eta^2) W_{\frac{ik^2}{4\tilde{m}}, \frac{1}{4}}(-i\tilde{m}\eta'^2). \quad (5.15)$$

It is reasonable to ask whether this vacuum state could potentially have observable effects. One way of approaching this problem is to calculate the response rate of a co-moving detector, such as the Unruh-Dewitt detector, when the field is in the S-J vacuum state. Even more ambitiously, one could (in principle) derive the S-J vacuum for a general scale-factor $a(\eta)$ (with reasonable boundary conditions), and study its back-reaction on the underlying geometry via the renormalized stress-energy-momentum tensor. These computations are fairly cumbersome and we defer a detailed treatment to future studies.

²¹ $\mathcal{W}\{W_{k,\mu}(z), W_{-k,\mu}(e^{\pm i\pi} z)\} = e^{\mp i k \pi}$ where \mathcal{W} is the Wronskian [22].

²²This is because $|g_k|^2 a(\eta)^2$ diverges quadratically in η , while $g_k^2 a(\eta)^2$ oscillates $\sim e^{-i\tilde{m}\eta^2}$.

In order to gain *some* intuition, however, we will compute $\langle SJ|\hat{\rho}|SJ\rangle \equiv \langle SJ|\hat{T}_0^0|SJ\rangle = a^{-2}\langle SJ|\hat{T}_{00}|SJ\rangle$, where the expectation value of the (un-renormalized) energy momentum tensor $\hat{T}_{\mu\nu}$ takes the form [21]:

$$\langle SJ|\hat{T}_{\mu\nu}|SJ\rangle = \int d^3\vec{k} T_{\mu\nu}(\mathcal{U}_k^{SJ}, \bar{\mathcal{U}}_k^{SJ}), \quad (5.16)$$

$$T_{\mu\nu}(\phi, \psi) \equiv \nabla_\mu \phi \nabla_\nu \psi - \frac{1}{2} g_{\mu\nu} [\nabla^\alpha \phi \nabla_\alpha \psi - m^2 \phi \psi]. \quad (5.17)$$

It can be checked that

$$T_{00}(\mathcal{U}_k^{SJ}, \bar{\mathcal{U}}_k^{SJ}) = \frac{1}{2} \left[|\partial_\eta \mathcal{U}_k^{SJ}|^2 + |\vec{\nabla} \mathcal{U}_k^{SJ}|^2 + m^2 a^2 |\mathcal{U}_k^{SJ}|^2 \right] \quad (5.18)$$

$$= \frac{1}{(2\pi)^3 2a^2} \left[\left| g'_k - \frac{a'}{a} g_k \right|^2 + (k^2 + m^2 a^2) |g_k|^2 \right], \quad (5.19)$$

where $'$ denotes differentiation with respect to η . Using the expression for $g_k(\eta)$ given by (5.12), it follows that

$$\langle SJ|\hat{\rho}|SJ\rangle = \int \frac{d^3\vec{p}}{(2\pi)^3} \sqrt{p^2 + m^2} (n_{SJ} + \frac{1}{2}), \quad (5.20)$$

where

$$n_{SJ} = \frac{1}{4} \frac{e^{-i\pi\lambda}}{\sqrt{\frac{p^2}{m^2} + 1}} \left\{ \left(\frac{p^2}{m^2} + 1 \right) |W_{\lambda,\mu}(z)|^2 - \frac{1}{z^2} \left| 2z \frac{dW_{\lambda,\mu}(z)}{dz} - \frac{3}{2} W_{\lambda,\mu}(z) \right|^2 \right\} - \frac{1}{2}. \quad (5.21)$$

The variables used above are defined as follows:

$$\vec{p} = \frac{\vec{k}}{a}, \quad H = \frac{a'}{a^2} = \frac{1}{\alpha\eta^2}, \quad z = i\tilde{m}\eta^2 = i\frac{m}{H}, \quad \lambda = \frac{-ik^2}{4\tilde{m}} = \frac{-ip^2}{4mH}. \quad (5.22)$$

As before $\mu = \frac{1}{4}$, and $p = |\vec{p}|$. In particular, H is the Hubble parameter and \vec{p} is the physical momentum of the Fourier mode with comoving wavenumber \vec{k} .

For a thermal bath of relativistic bosons at temperature T , the energy density takes the form $\rho = \int \frac{d^3\vec{p}}{(2\pi)^3} E n_{BE}$, where $E = \sqrt{p^2 + m^2}$ and $n_{BE} = \frac{1}{e^{E/T} - 1}$ is the Bose-Einstein distribution. This relation can be inverted to get $T = \frac{E}{\ln(1 + \frac{1}{n_{BE}})}$. In order to see how “close to thermal” our state is, we similarly define the “effective temperature” of a mode as

$$T_{SJ}(p) = \frac{\sqrt{p^2 + m^2}}{\ln\left(1 + \frac{1}{n_{SJ}}\right)}. \quad (5.23)$$

The more constant T_{SJ} is as a function of p , the closer the distribution n_{SJ} is to being thermal. Here we define n_{SJ} to include only excitations *above* the state $|GS\rangle$ that minimizes the Hamiltonian at a particular instant of time, and for which $\langle GS|\hat{\rho}|GS\rangle = \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{2} \sqrt{p^2 + m^2}$.

Then $\langle SJ|\hat{\rho}|SJ\rangle - \langle GS|\hat{\rho}|GS\rangle = \int \frac{d^3\vec{p}}{(2\pi)^3} \sqrt{p^2 + m^2} n_{SJ}$.

Figure 1 shows the behaviour of T_{SJ} for different ratios of $\frac{m}{H}$ and $\frac{p}{m}$. It is evident that

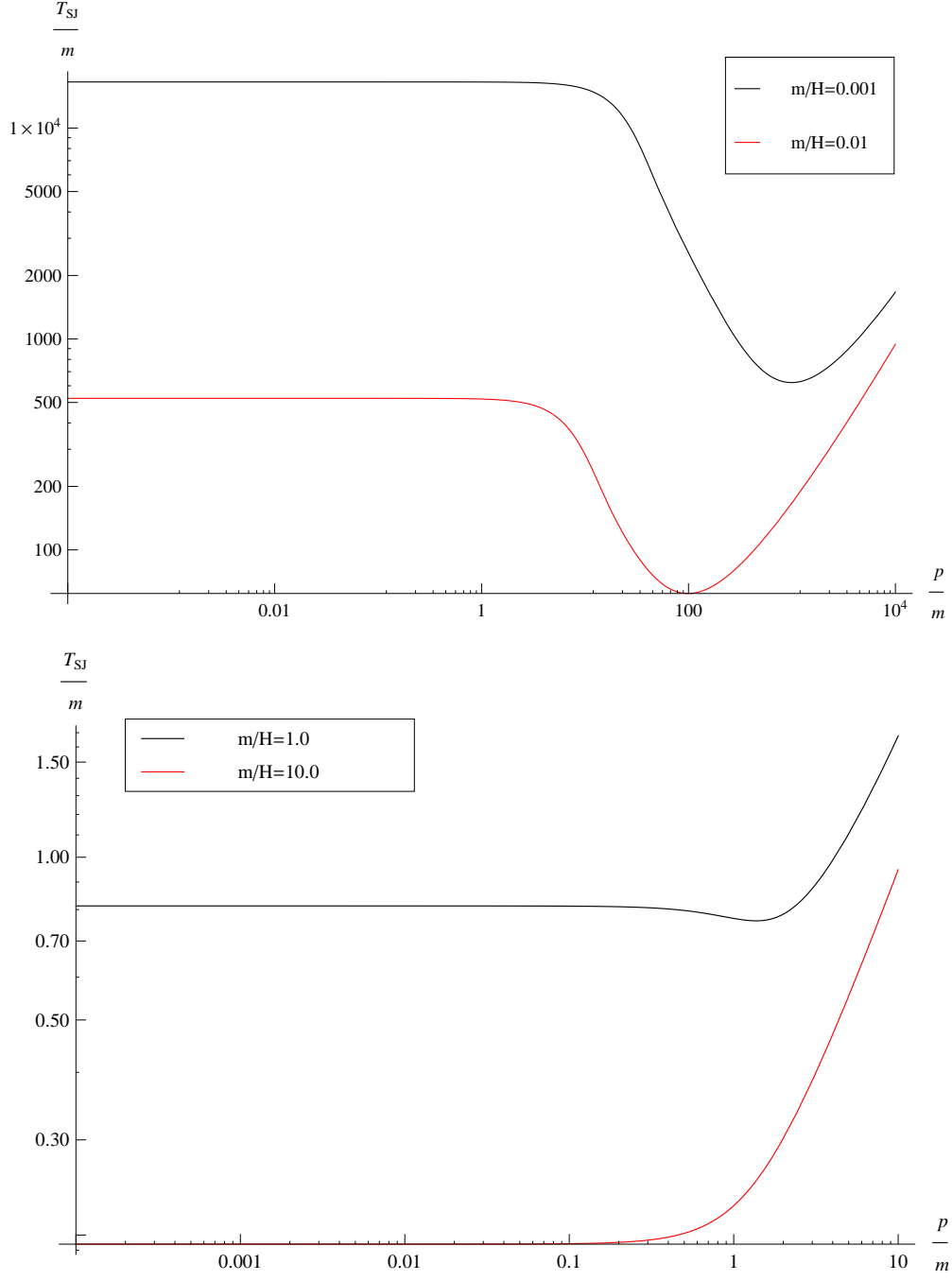


Figure 1. Behaviour of T_{SJ} for different ratios of $\frac{m}{H}$ and $\frac{p}{H}$.

the long wavelength modes are in fact at a constant effective temperature. For example, in the regime where $m \ll H$ and $p \ll \sqrt{mH}$, the Whittaker function has a simple asymptotic expansion

$$W_{\lambda, \frac{1}{4}}(z) \longrightarrow \frac{\sqrt{\pi}}{\Gamma(\frac{3}{4} - \lambda)} z^{\frac{1}{4}} - \frac{2\sqrt{\pi}}{\Gamma(\frac{1}{4} - \lambda)} z^{3/2} + \mathcal{O}(z^{5/4}), \quad (5.24)$$

using which it can be shown that

$$T_{SJ}(p) \sim \frac{\pi}{4|\Gamma(3/4)|^2} \frac{H^{3/2}}{\sqrt{m}}. \quad (5.25)$$

This result suggests that there are correlations on super-horizon scales. It is noteworthy that these correlations have appeared without the help of any previous epoch of accelerated expansion. Potentially, they could therefore open up a new perspective on the question of primordial fluctuations and on the related puzzle sometimes called “horizon problem”.

6 Causality and the S-J Vacuum

Like other vacua, the S-J vacuum is defined globally, and it depends on both the causal past *and* future of the spacetime. Consider for example a spacetime which is first static, then expands for a short time, then goes back to being static again.²³ In light of the inherent time-reversal symmetry of the conditions defining our vacuum-state, it is clear that this state can agree neither with the early-time vacuum (the state of minimum energy at that time), nor with the late-time vacuum. Rather, it must strike some sort of “compromise” between them.

In the present section we will illustrate this behavior with a simple example, but before doing so, we would like to dwell for a bit on the question of whether one should interpret this type of dependence on the future as a failure of causality. By construction, our definition of the vacuum depends on the full spacetime geometry. That it thereby fails to be what John Bell called “locally causal” is no surprise because, as is well understood by now, any reasonable quantum state must incorporate nonlocal correlations and entanglement. Certainly the Minkowski vacuum does so. But does this type of nonlocality also imply genuine acausality?

The prior question that begs for an answer here is what is meant by acausality in the context of quantum field theory, considering also that quantum field theory must ultimately find its place within a theory of full quantum gravity. If we remain within the “operationalist” framework of external agents, “measurements” and state-vector collapse, then causality (in the sense of relativistic causality) reduces to the impossibility of superluminal signalling. In this sense, there is no question of acausality as long as the twin conditions of spacelike commutativity and hyperbolicity of the field equations are respected, which by construction they are in the field theory we are working with in this paper.²⁴ On the other hand, if we try to adopt a more “objective” framework which dispenses with external agents, then we seem to be left without any clear definition of relativistic causality at all. That is, we lack an *intrinsic* criterion which could decide whether or not physical influences are propagating outside the light cone or “into the past”. But without such a criterion, the meaning of relativistic causality in general is called into question.

²³Of course, such a spacetime is not necessarily a solution to the Einstein equations.

²⁴The theory of [1] retains spacelike commutativity, but hyperbolicity becomes, together with the notion of field-equation itself, approximate at best.

A further observation also seems relevant here, even if it does not turn out to be decisive. Namely, the assumption we have made of a fixed, non-dynamical spacetime is already “anticausal” in a certain sense. In a full quantum gravity theory the future geometry must evolve together with, and in mutual dependence on the future matter-field. Hence, any attempt to specify the geometry in advance amounts to imposing a future boundary condition on the combined system of metric plus scalar field. Given this, it would not be surprising if a correct semiclassical treatment of the scalar were also to involve some degree of “dependence on the future”.

The specific model we will consider is a 1+1 dimensional FLRW universe with metric $ds^2 = C(\eta)(d\eta^2 - dx^2)$, where $C(\eta) = A + B \tanh(\rho\eta)$. In the infinite past $C(\eta) \rightarrow A - B$ and in the infinite future $C(\eta) \rightarrow A + B$. It is known that there are normalized modes $u_k^{in}(\eta, x)$ that behave like positive frequency Minkowski-space modes in the remote past $(\eta, t \rightarrow -\infty)$:²⁵

$$\begin{aligned} u_k^{in}(\eta, x) &= \frac{1}{\sqrt{4\pi\omega_k^{in}}} e^{ikx - i\omega_k^+ \eta - (i\omega_k^-/\rho) \ln[2 \cosh(\rho\eta)]} \\ &\times {}_2F_1(1 + (i\omega_k^-/\rho), i\omega_k^-/\rho; 1 - (i\omega_k^{in}/\rho); \frac{1}{2}(1 + \tanh(\rho\eta))) \\ &\xrightarrow{\eta \rightarrow -\infty} \frac{1}{\sqrt{4\pi\omega_k^{in}}} e^{ikx - i\omega_k^{in} \eta}, \end{aligned} \quad (6.1)$$

where ${}_2F_1$ is the ordinary hypergeometric function and

$$\begin{aligned} \omega_k^{in} &= [k^2 + m^2(A - B)]^{1/2} \\ \omega_k^{out} &= [k^2 + m^2(A + B)]^{1/2} \\ \omega_k^\pm &= \frac{1}{2}(\omega_k^{out} \pm \omega_k^{in}). \end{aligned} \quad (6.2)$$

Similarly, there are normalized modes $u_k^{out}(\eta, x)$ that behave like the positive frequency Minkowski-space modes in the remote future: $(\eta, t \rightarrow \infty)$

$$\begin{aligned} u_k^{out}(\eta, x) &= \frac{1}{\sqrt{4\pi\omega_k^{out}}} e^{ikx - i\omega_k^+ \eta - (i\omega_k^-/\rho) \ln[2 \cosh(\rho\eta)]} \\ &\times {}_2F_1(1 + (i\omega_k^-/\rho), i\omega_k^-/\rho; 1 + (i\omega_k^{out}/\rho); \frac{1}{2}(1 - \tanh(\rho\eta))) \\ &\xrightarrow{\eta \rightarrow \infty} \frac{1}{\sqrt{4\pi\omega_k^{out}}} e^{ikx - i\omega_k^{out} \eta}. \end{aligned} \quad (6.3)$$

The in and out modes are related to each other by the following Bogolubov transformation

$$u_k^{in}(\eta, x) = \alpha_k u_k^{out}(\eta, x) + \beta_k \bar{u}_{-k}^{out}, \quad (6.4)$$

²⁵See section 3.4 of [21].

where

$$\alpha_k = \left(\frac{\omega_k^{out}}{\omega_k^{in}}\right)^{1/2} \frac{\Gamma(1 - (i\omega_k^{in}/\rho))\Gamma(-(i\omega_k^{out}/\rho))}{\Gamma(1 - (i\omega_k^+/\rho))\Gamma(-i\omega_k^+/\rho)} \quad (6.5)$$

$$\beta_k = \left(\frac{\omega_k^{out}}{\omega_k^{in}}\right)^{1/2} \frac{\Gamma(1 - (i\omega_k^{in}/\rho))\Gamma(-(i\omega_k^{out}/\rho))}{\Gamma(1 + (i\omega_k^-/\rho))\Gamma(i\omega_k^-/\rho)}. \quad (6.6)$$

The modes u_k^{in} and u_k^{out} define vacuum states at early and late times, respectively. If the system is at first ($\eta \rightarrow -\infty$) in the in-vacuum state, i.e. the no particle state, it will have $|\beta_k|^2$ particles of momentum k with respect to the out-vacuum after the expansion ($\eta \rightarrow \infty$). The S-J vacuum has a different nature, simply because the vacuum state in the $\eta \rightarrow -\infty$ region depends on what happens in the infinite future (and vice-versa). We can find the S-J vacuum by substituting the modefunctions $u_k^{in}(\eta, x)$ in formula (3.15). Defining $g_k(\eta)$ through $u_k^{in}(\eta, x) = \frac{e^{ikx}}{\sqrt{2\pi}}g_k(\eta)$, it can be easily verified that

$$\langle u_k^{in}, u_{k'}^{in} \rangle = \delta(k - k') \int_{-\Lambda}^{\Lambda} \bar{g}_k(\eta) g_{k'}(\eta) C(\eta) d\eta \quad (6.7)$$

$$\langle \bar{u}_k^{in}, u_{k'}^{in} \rangle = \delta(k + k') \int_{-\Lambda}^{\Lambda} g_k(\eta) g_{k'}(\eta) C(\eta) d\eta, \quad (6.8)$$

where as usual, we have regulated the integrals with a cutoff Λ . The asymptotic behaviour of $g_k(\eta)$ is given by

$$g_k(\eta) \xrightarrow{\eta \rightarrow -\infty} \frac{1}{\sqrt{2\omega_k^{in}}} e^{-i\omega_k^{in}\eta} \quad (6.9)$$

$$g_k(\eta) \xrightarrow{\eta \rightarrow \infty} \frac{1}{\sqrt{2\omega_k^{out}}} (\alpha_k e^{-i\omega_k^{out}\eta} + \beta_k e^{i\omega_k^{out}\eta}). \quad (6.10)$$

Using these expressions, it can be checked that

$$\lim_{\Lambda \rightarrow \infty} \frac{\langle \bar{u}_{-k}^{in}, u_k^{in} \rangle}{\langle u_k^{in}, u_k^{in} \rangle} \rightarrow \frac{2\alpha_k \beta_k}{|\alpha_k|^2 + |\beta_k|^2 + \frac{\omega_{out}}{\omega_{in}} \frac{A-B}{A+B}} \equiv \gamma_k, \quad (6.11)$$

from which the S-J vacuum can be computed using (3.15):

$$u_k^{SJ}(\eta, x) = \mu_k u_k^{in}(\eta, x) + \xi_k \bar{u}_{-k}^{in}(\eta, x) \quad (6.12)$$

$$\mu_k = \frac{1}{\sqrt{2}} \left[1 + \frac{1}{\sqrt{1 - |\gamma_k|^2}} \right]^{\frac{1}{2}} \quad (6.13)$$

$$\xi_k = -\frac{1}{\sqrt{2}} \left[-1 + \frac{1}{\sqrt{1 - |\gamma_k|^2}} \right]^{\frac{1}{2}} e^{i\text{Arg}(\alpha_k \beta_k)}. \quad (6.14)$$

Fig. 2 shows the difference between the S-J and “in” vacua for a specific set of parameters and frequencies. As one would expect, this deviation is only significant for low-frequency modes, which are more sensitive to the rate of expansion ρ .

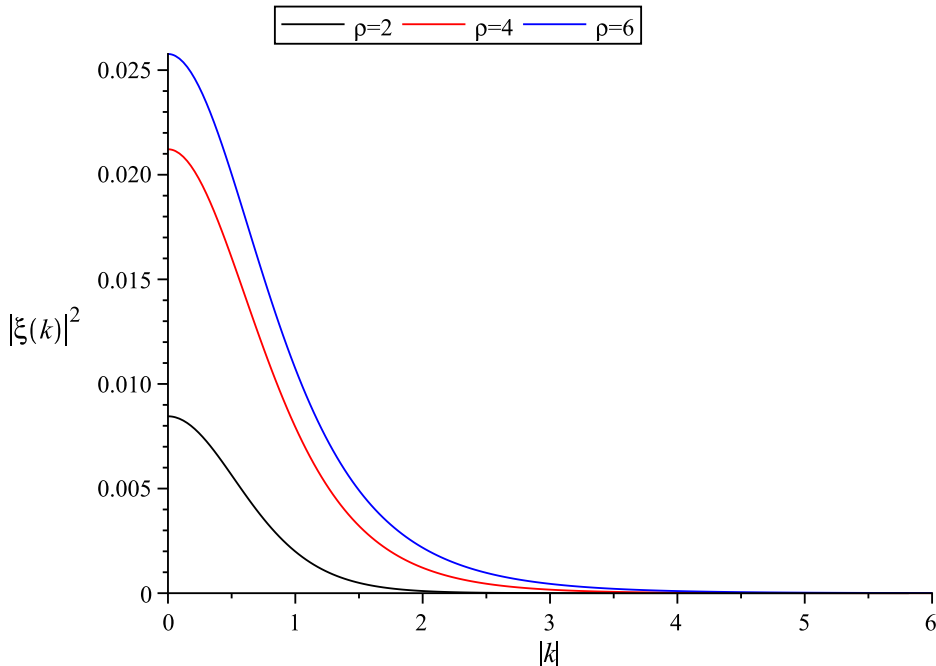


Figure 2. The deviation between the S-J and “in” vacua, as measured by $|\xi(k)|^2$, defined in (6.14). The parameters used here are $A = 2.0$, $B = 1.0$, and $m = 1.0$.

7 Conclusions and Discussions

We have defined a distinguished vacuum for a free quantum field in a globally hyperbolic region of an arbitrarily curved spacetime. This “S-J” state is well-defined for all compact regions and for a large class of noncompact ones.

We have shown that for static spacetimes, our vacuum coincides with the usual ground state. We have also computed it explicitly for a scalar field of mass $m \geq 0$ in a radiation-filled, spatially flat, homogeneous and isotropic cosmos. In that connection we also computed an “effective temperature that can be defined for the super-horizon modes of the massive field. The correlations found thereby could open up a new perspective on the question of primordial fluctuations and the so-called “horizon problem”.

A peculiar aspect of our prescription is its temporal non-locality. We demonstrated this feature by the example of a spacetime which sandwiches a region with curvature in-between flat initial and final regions, but we did not explore its phenomenological implications any further. In a parallel effort [18], we have also applied our prescription to de Sitter space, obtaining results for both the full spacetime and for the so-called Poincaré patch. The vacua obtained thereby differ from the Euclidean (Bunch-Davies) vacuum below a certain mass threshold, with potentially interesting phenomenology.

A question that we have not addressed in this paper is whether, or in which circumstances, the S-J vacuum obeys the so-called Hadamard condition. Since this work was completed, some results have appeared [16] showing that the answer is yes in some cases and no in others. We hope to return to this and related matters elsewhere.

Acknowledgments

We would like to thank Yasaman K. Yazdi, Fay Dowker, Michel Buck, and Bill Unruh for useful discussions and comments throughout the course of this project. We are supported by the University of Waterloo and the Perimeter Institute for Theoretical Physics. Research at the Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research & Innovation.

A Comments and Calculations

A.1 When is $i\Delta$ self-adjoint?

We saw in the main text that $i\Delta$ is self-adjoint and has a complete set of normalizable eigenvectors when $i\Delta(x, y) \in L^2(M \times M)$. In 3+1 dimensions, this Hilbert-Schmidt condition is not satisfied, however, because the retarded Green's function is a distribution, and not a function. Nonetheless, it is still possible to show that $i\Delta$ is self-adjoint when M is bounded. Here, we will show that this is indeed the case for a bounded region M of Minkowski space, and will argue that this conclusion should continue to hold in all curved spacetimes.

In 3 + 1 dimensional Minkowski space (\mathbb{M}^4), we have $G_R(x, y) = G(x - y)$ with $G(x)$ given by

$$G(x) = \theta(t)\theta(s^2)\left(\frac{\delta(s^2)}{2\pi} - \frac{m}{4\pi s}J_1(ms)\right), \quad (\text{A.1})$$

where $s = \sqrt{t^2 - \vec{x}^2}$ for $t^2 \geq \vec{x}^2$, and $s = -i\sqrt{\vec{x}^2 - t^2}$ for $t^2 \leq \vec{x}^2$. J_1 is a Bessel function of the first kind and m is the mass of the scalar field. We will now show that $i\Delta$ is a bounded operator on $C_0^\infty(M)$, i.e. we will prove that there exists $N > 0$ such that $\|i\Delta f\| \leq N\|f\|$ for all $f \in C_0^\infty(M)$, where $\|\cdot\|$ is the $L^2(\mathbb{M}^4)$ norm.²⁶ Let $G_R f = \int_M G_R(x, y)f(y) dV_y$ and $G_A f = \int_M G_A(x, y)f(y) dV_y$ denote the retarded and advanced solutions of the Klein-Gordon equation with source f , respectively. It is enough to show that G_R and G_A are bounded because $\|i\Delta f\| = \|iG_R f - iG_A f\| \leq \|G_R f\| + \|G_A f\|$.

Let $A(x) = \theta(t)\theta(s^2)\frac{\delta(s^2)}{2\pi}$ and $B(x) = \theta(t)\theta(s^2)\frac{m}{4\pi s}J_1(ms)$ so that $G(x) = A(x) - B(x)$. Then: $\|G_R f\| = \|A f - B f\| \leq \|A f\| + \|B f\| \leq \|A f\| + \frac{m^2 V}{8\pi}\|f\|$, where V is the total spacetime volume.²⁷ Therefore, G_R is bounded if and only if A is bounded. Consider a bounded region of \mathbb{M}^4 where $\vec{x} \in [\vec{x}_{min}, \vec{x}_{max}]$ and $t \in [t_{min}, t_{max}]$, and a smooth function of compact support $f(t, \vec{x})$ on this region. It can be shown that:

$$(A f)(t, \vec{x}) = \int_{\vec{x}_{min}}^{\vec{x}_{max}} \frac{f(t - |\vec{x} - \vec{y}|, \vec{y})}{4\pi|\vec{x} - \vec{y}|} d^3\vec{y}. \quad (\text{A.2})$$

It then follows from the Cauchy-Schwarz inequality that

$$|(A f)(t, \vec{x})|^2 \leq \frac{1}{16\pi^2} \left\{ \int_{\vec{x}_{min}}^{\vec{x}_{max}} \frac{d^3\vec{y}}{|\vec{x} - \vec{y}|^2} \right\} \times \left\{ \int_{\vec{x}_{min}}^{\vec{x}_{max}} |f(t - |\vec{x} - \vec{y}|, \vec{y})|^2 d^3\vec{y} \right\}. \quad (\text{A.3})$$

²⁶ $\|\psi\| = \sqrt{\langle \psi, \psi \rangle}$.

²⁷ B is Hilbert-Schmidt because $\int_M \int_M |B(x - y)|^2 dV_x dV_y \leq (\frac{m^2 V}{8\pi})^2$, since $\frac{J_1(ms)}{s}$ peaks at $\frac{m}{2}$.

Also,

$$\int_{\vec{x}_{min}}^{\vec{x}_{max}} \frac{d^3\vec{y}}{|\vec{x} - \vec{y}|^2} = \int_{\vec{x}_{min} + \vec{x}}^{\vec{x}_{max} + \vec{x}} \frac{d^3\vec{y}}{|\vec{y}|^2} \leq \int_{2\vec{x}_{min}}^{2\vec{x}_{max}} \frac{d^3\vec{y}}{|\vec{y}|^2} \equiv C_0, \quad (\text{A.4})$$

where C_0 is some finite positive number. So we have that

$$|(Af)(t, \vec{x})|^2 \leq \frac{C_0}{16\pi^2} \left\{ \int_{\vec{x}_{min}}^{\vec{x}_{max}} |f(t - |\vec{x} - \vec{y}|, \vec{y})|^2 d^3\vec{y} \right\}. \quad (\text{A.5})$$

It then follows that

$$\|Af\|^2 = \int_{\vec{x}_{min}}^{\vec{x}_{max}} d^3\vec{x} \int_{t_{min}}^{t_{max}} dt |(Af)(t, \vec{x})|^2 \quad (\text{A.6})$$

$$\leq \frac{C_0}{16\pi^2} \int_{\vec{x}_{min}}^{\vec{x}_{max}} d^3\vec{x} \int_{\vec{x}_{min}}^{\vec{x}_{max}} d^3\vec{y} \int_{t_{min}}^{t_{max}} dt |f(t - |\vec{x} - \vec{y}|, \vec{y})|^2 \quad (\text{A.7})$$

$$= \frac{C_0}{16\pi^2} \int_{\vec{x}_{min}}^{\vec{x}_{max}} d^3\vec{x} \int_{\vec{x}_{min}}^{\vec{x}_{max}} d^3\vec{y} \int_{t_{min} + |\vec{x} - \vec{y}|}^{t_{max} + |\vec{x} - \vec{y}|} dt |f(t, \vec{y})|^2. \quad (\text{A.8})$$

Since $f(t, \vec{y})$ vanishes for $t > t_{max}$: $\int_{t_{min} + |\vec{x} - \vec{y}|}^{t_{max} + |\vec{x} - \vec{y}|} dt |f(t, \vec{y})|^2 \leq \int_{t_{min}}^{t_{max}} dt |f(t, \vec{y})|^2$, from which it follows that

$$\|Af\|^2 \leq \frac{C_0 V_s}{16\pi^2} \|f\|^2, \quad (\text{A.9})$$

where $V_s = \int_{\vec{x}_{min}}^{\vec{x}_{max}} d^3\vec{x}$ is the enclosed spatial volume. A similar analysis goes through for the advanced solution $G_A f$ which results in the same bound. At long last:

$$\|i\Delta f\| \leq \frac{V}{4\pi} \left(m^2 + 2\sqrt{\frac{C_0 V_s}{V^2}} \right) \|f\|. \quad (\text{A.10})$$

Because $i\Delta$ is bounded on $C_0^\infty(M)$ and $C_0^\infty(M)$ is a dense subspace of $L^2(M)$, $i\Delta$ can be uniquely extended to $L^2(M)$ as a bounded operator [15]. It then follows that $i\Delta$ is self-adjoint because it is Hermitian on all of $L^2(M)$ [15].

It should be clear that the boundedness of $i\Delta$ has everything to do with the singularity structure of $i\Delta(x, y)$ (once we restrict ourselves to bounded spacetimes). It is not terribly unrealistic to assume that this singularity structure remains (more or less) the same in curved spacetimes. This is certainly true in the coincidence limit, if the equivalence principle is respected. Based on these arguments, we assume in this paper that $i\Delta$ is a self-adjoint operator on $L^2(M)$, for all bounded globally-hyperbolic spacetimes M .

A.2 The S-J Vacuum and the Simple Harmonic Oscillator

This simple example illustrates the technical difficulties one faces when diagonalizing $i\Delta$, and how they can be resolved. Consider a simple harmonic oscillator with unit mass and frequency ω , whose position $q(t)$ satisfies $(\frac{d^2}{dt^2} + \omega^2)q(t) = 0$. The associated retarded Green's function satisfies $(\frac{d^2}{dt^2} + \omega^2)G_R(t, t') = -\delta(t - t')$, with $G_R(t, t') = 0$ for $t < t'$. The solution to this equation is $G_R(t, t') = -\theta(t - t') \frac{\sin[\omega(t - t')]}{\omega}$, which in turn gives

$$\Delta(t, t') = G_R(t, t') - G_R(t', t) = -\frac{\sin[\omega(t - t')]}{\omega} \quad (\text{A.11})$$

$$= \frac{1}{2i\omega} [e^{-i\omega(t - t')} - e^{i\omega(t - t')}] \quad (\text{A.12})$$

Taking $t \in (-\infty, \infty)$, it may be verified that, ²⁸ formally,

$$i\Delta e^{\pm i\omega t} = \frac{\mp \delta(0)}{2\omega} e^{\pm i\omega t}. \quad (\text{A.13})$$

Keeping the $\delta(0)$'s around, we see that $T_{\mp}(t) = \frac{e^{\pm i\omega t}}{\sqrt{\delta(0)}}$ are orthonormal eigenfunctions of $i\Delta$ with eigenvalues $\lambda_{\pm} = \pm \frac{\delta(0)}{2\omega}$. According to our prescription, the resulting positive frequency modefunction $\mathcal{U}^{SJ}(t) = \sqrt{\lambda_+} T_+(t) = \frac{e^{-i\omega t}}{\sqrt{2\omega}}$, which is completely well defined and gives the right vacuum state: the state $|0\rangle$ annihilated by \hat{a} which multiplies $\mathcal{U}^{SJ}(t)$ in the position operator expansion $\hat{q}(t) = \mathcal{U}^{SJ}(t)\hat{a} + \bar{\mathcal{U}}^{SJ}(t)\hat{a}^\dagger$ is in fact the minimum energy state of the Hamiltonian. Thus the infinities appearing in the spectrum of $i\Delta$ end up being harmless. In other words, the S-J vacuum state should not depend on how $\delta(0)$ is regularized.

One regularization scheme, for example, is to first restrict to $t \in [-T, T]$, diagonalize $i\Delta$, and then take the limit $T \rightarrow \infty$ once the spectrum of $i\Delta$ has been computed. Let $u(t) = \frac{e^{-i\omega t}}{\sqrt{2\omega}}$ so that $i\Delta(t, t') = u(t)\bar{u}(t') - u(t')\bar{u}(t)$. Finding the spectrum of $i\Delta$ in this case is similar to diagonalizing a two-by-two matrix. It may be confirmed that there are two eigenvalues λ and $-\lambda$ with corresponding eigenfunctions T_+ and \bar{T}_+ , where

$$T_+ = \sqrt{\frac{\langle u, u \rangle + \lambda}{2\lambda^2}} u(t) - \sqrt{\frac{\langle u, u \rangle - \lambda}{2\lambda^2}} e^{-i\arg(\langle u, \bar{u} \rangle)} \bar{u}(t) \quad (\text{A.14})$$

$$\lambda = \sqrt{\langle u, u \rangle^2 - |\langle u, \bar{u} \rangle|^2}. \quad (\text{A.15})$$

These expressions are completely well-defined because all the inner products are finite. The corresponding positive frequency modefunction as dictated by our prescription is then

$$\mathcal{U}^{SJ}(t) = \sqrt{\lambda} T_+(t) = \frac{1}{\sqrt{2}} \left\{ \sqrt{1 + \frac{1}{\sqrt{1 - \frac{|\langle u, \bar{u} \rangle|^2}{\langle u, u \rangle^2}}}} u(t) - \sqrt{\frac{1}{\sqrt{1 - \frac{|\langle u, \bar{u} \rangle|^2}{\langle u, u \rangle^2}}}} - 1 e^{-i\arg(\langle u, \bar{u} \rangle)} \bar{u}(t) \right\}. \quad (\text{A.16})$$

In the limit $T \rightarrow \infty$, the ratio $\frac{|\langle u, \bar{u} \rangle|^2}{\langle u, u \rangle^2} \rightarrow 0$ and we recover $\mathcal{U}^{SJ}(t) = \frac{e^{-i\omega t}}{\sqrt{2\omega}}$.

A.3 Equation (3.4) as an equality between distributions

Let us show that the right and left hand sides of (3.4) are equal if they are integrated against a smooth test function $f \in C_0^\infty(M)$.

Since $\Delta f \in \mathbf{S}$ for any $f \in C_0^\infty(M)$, we can expand out $i\Delta f$ in terms of the u_k : $i\Delta f(x) = \sum_k \alpha_k u_k(x) + \beta_k \bar{u}_k(x)$, where α_k 's and β_k 's are constants. It can be verified that $\alpha_k = (u_k, i\Delta f)_{KG}$ and $\beta_k = -(\bar{u}_k, i\Delta f)_{KG}$. Then,

$$i\Delta f(x) = \sum_k (u_k, i\Delta f)_{KG} u_k(x) - (\bar{u}_k, i\Delta f)_{KG} \bar{u}_k(x) \quad (\text{A.17})$$

$$= \sum_k \langle u_k, f \rangle u_k(x) - \langle \bar{u}_k, f \rangle \bar{u}_k(x) \quad (\text{A.18})$$

$$= \int_M \left[\sum_k u_k(x) \bar{u}_k(y) - \bar{u}_k(x) u_k(y) \right] f(y) dV_y. \quad (\text{A.19})$$

²⁸Here, as always, $(\Delta f)(t) = \int \Delta(t, t') f(t') dt'$.

A.4 Another Slant on the SJ Prescription

The following simple but useful result helps put our construction in context.

Theorem A.1. Let $\mathbf{S}_{\mathbb{C}}$ denote the set of all complex solutions of the Klein-Gordon equation. Assume there are functions $u_i(x) \in \mathbf{S}_{\mathbb{C}}$, which together with their complex conjugates span $\mathbf{S}_{\mathbb{C}}$ and satisfy

$$(u_n, u_m)_{KG} = \delta_{nm}, \quad (\bar{u}_n, \bar{u}_m)_{KG} = -\delta_{nm}, \quad (u_n, \bar{u}_m)_{KG} = 0 \quad (\text{A.20})$$

$$\frac{\langle u_n, u_m \rangle}{\langle u_n, u_n \rangle} = \delta_{nm}, \quad \langle u_n, \bar{u}_m \rangle = 0, \quad (\text{A.21})$$

where $\langle, \rangle: \mathbf{S}_{\mathbb{C}} \times \mathbf{S}_{\mathbb{C}} \rightarrow \mathbb{C}$ is the L^2 inner product defined in (3.1). (See Section 2 for the definition of $(,)_{KG}$.) Then, any other functions $g_n(x) \in \mathbf{S}_{\mathbb{C}}$ that also satisfy these conditions are unitarily-related to the $u_i(x)$ (i.e. $g_n(x) = \sum_k \alpha_{nk} u_k(x)$ where $\sum_k \bar{\alpha}_{nk} \alpha_{mk} = \delta_{nm}$).

Proof. Suppose both $u_n(x)$ and $g_n(x)$ satisfy the above conditions and let

$$g_n(x) = \sum_k \alpha_{nk} u_k(x) + \beta_{nk} \bar{u}_k(x). \quad (\text{A.22})$$

Requiring $(g_n, g_m)_{KG} = \delta_{nm}$ implies

$$\sum_k \bar{\alpha}_{nk} \alpha_{mk} - \bar{\beta}_{nk} \beta_{mk} = \delta_{nm}. \quad (\text{A.23})$$

Also, it follows from (A.22) and (A.20) that

$$\alpha_{nk} = (u_k, g_n)_{KG}, \quad \beta_{nk} = -(\bar{u}_k, g_n)_{KG}. \quad (\text{A.24})$$

Given that $\sum_k |g_k\rangle_{KG} \langle g_k| - |\bar{g}_k\rangle_{KG} \langle \bar{g}_k|$ is the identity operator on $\mathbf{S}_{\mathbb{C}}$, it follows from (A.24) that

$$\begin{aligned} u_n(x) &= \sum_k (g_k, u_n)_{KG} g_k(x) - (\bar{g}_k, u_n)_{KG} \bar{g}_k(x) \\ &= \sum_k \bar{\alpha}_{kn} g_k(x) - \beta_{kn} \bar{g}_k(x). \end{aligned} \quad (\text{A.25})$$

Using (A.22) and (A.21), it can also be verified that

$$\beta_{nk} = \frac{\langle \bar{u}_k, g_n \rangle}{\langle u_k, u_k \rangle}. \quad (\text{A.26})$$

Similarly, (A.25) and the L^2 orthogonality conditions for the $g_i(x)$ s (which amounts to replacing all u 's with g 's in (A.21)) imply

$$\beta_{nk} = -\frac{\langle \bar{g}_n, u_k \rangle}{\langle g_n, g_n \rangle}. \quad (\text{A.27})$$

Given that $\langle \bar{u}_k, g_n \rangle = \langle \bar{g}_n, u_k \rangle$, $\langle u_k, u_k \rangle > 0$, and $\langle g_n, g_n \rangle > 0$, it follows from (A.26) and (A.27) that $\langle \bar{u}_k, g_n \rangle = 0$. This in turn implies that $\beta_{nk} = 0$, which together with (A.23) proves the theorem. \square

Note that this proof is valid for any Hermitian inner product \langle, \rangle on $\mathbf{S}_{\mathbb{C}}$ which enjoys the additional property $\langle \bar{f}, g \rangle = \langle \bar{g}, f \rangle$.

We have shown here that requiring (A.21), in addition to the usual quantization conditions, is enough to define a distinguished vacuum state. That these modes are eigenfunctions of $i\Delta$ ($(i\Delta u_n)(x) = \langle u_n, u_n \rangle u_n(x)$), follows from (3.4).

B Useful Theorems

Theorem B.1. ²⁹ Let (M, g_{ab}) be a globally hyperbolic spacetime with smooth, spacelike Cauchy surface Σ . Then the Klein-Gordon equation (2.1) has a well posed initial value formulation in the following sense: Given any pair of smooth (C^∞) functions $(\phi_0, \dot{\phi}_0)$ on Σ , there exists a unique solution ϕ to (2.1), defined on all of M , such that on Σ we have $\phi = \phi_0$ and $n^a \nabla_a \phi = \dot{\phi}_0$, where n^a denotes the unit (future-directed) normal to Σ . Furthermore, for any closed subset $S \subset \Sigma$, the solution, ϕ , restricted to $D(S)$ depends only upon the initial data on S . In addition, ϕ is smooth and varies continuously with the initial data.

The above theorem continues to hold if a (fixed, smooth) “source term” f is inserted on the right hand side of (2.1). It then follows from the domain of dependence feature of the theorem that there exist unique advanced and retarded solutions to the Klein-Gordon equation with source, whence in a globally hyperbolic spacetime there exist unique advanced and retarded Green’s functions for the Klein-Gordon equation.

Theorem B.2. ³⁰ $\Delta : C_0^\infty \rightarrow \mathbf{S}$ satisfies $\int_M \phi f \, dV = \Omega(\Delta f, \phi)$ for all $f \in C_0^\infty$ and $\phi \in \mathbf{S}$.

Proof. Let’s first give a simple proof for 1+1 dimensional Minkowski space and then generalize the arguments to globally hyperbolic spacetimes of higher spatial dimension. Pick $t_1, t_2 \in \mathbb{R}$ such that $f = 0$ for $t \notin [t_1, t_2]$. Then:

$$\int_M \phi f \, dV = \int_{t_1}^{t_2} dt \int_{-\infty}^{\infty} \phi (\partial_t^2 - \partial_x^2 + m^2)(Af) \, dx, \quad (\text{B.1})$$

where Af is the advanced solution of the Klein-Gordon equation with source f . Now integrating by parts twice yields

$$\int_{t_1}^{t_2} \phi \partial_t^2(Af) \, dt = \phi \partial_t(Af)|_{t_1}^{t_2} - \partial_t \phi(Af)|_{t_1}^{t_2} + \int_{t_1}^{t_2} (Af) \partial_t^2 \phi \, dt \quad (\text{B.2})$$

$$= \partial_t \phi(Af)|_{t_1} - \phi \partial_t(Af)|_{t_1} + \int_{t_1}^{t_2} (Af)(\partial_x^2 - m^2)\phi \, dt, \quad (\text{B.3})$$

where we have used the Klein-Gordon equation and that Af vanishes outside the causal past of the support of f . Substituting this back into (B.1), we can further simplify by noting

²⁹This is an exact restatement of Theorem 4.1.2 of [8], which we have included for the convenience of the reader. The symbol ‘ $D(S)$ ’ denotes the so-called domain of dependence of S .

³⁰This is a generalization of Lemma 3.2.1 of [8].

that $\int_{-\infty}^{\infty} [(Af)\partial_x^2\phi - \phi\partial_x^2(Af)] dx = 0$. (Because Af induces initial data of compact support on all equal-time spatial slices, all surface terms vanish.) Using this we find:

$$\int_M \phi f dV = \int_{-\infty}^{\infty} [\partial_t\phi(Af)|_{t_1} - \phi\partial_t(Af)|_{t_1}] dx = \Omega(\Delta f, \phi), \quad (\text{B.4})$$

since Rf (the retarded solution of the Klein-Gordon equation with source f) vanishes at $t = t_1$. To generalize the proof, we write

$$\begin{aligned} \int_M \phi f dV &= \int_M \phi(\nabla_\mu\nabla^\mu + m^2)(Af) dV \\ &= \int_M \nabla_\mu\{\phi\nabla^\mu(Af)\} dV - \int_M \nabla_\mu\phi\nabla^\mu(Af) dV + \int_M m^2\phi(Af) dV \\ &= \int_M \nabla_\mu\{\phi\nabla^\mu(Af) - (Af)\nabla^\mu\phi\} dV + \int_M (\nabla_\mu\nabla^\mu + m^2)\phi(Af) dV \\ &= \int_{\partial M} \{\phi n^\mu\nabla_\mu(Af) - (Af)n^\mu\nabla_\mu\phi\}\sqrt{-h} d^3x, \end{aligned} \quad (\text{B.5})$$

where we have used the Stoke's theorem in the last line. Here n^μ is the unit normal to the boundary ∂M and h is the determinant of the induced metric on the boundary. Again, because Af induces initial data of compact support on all equal-time spatial slices and Af vanishes outside the causal past of the support of f :

$$\begin{aligned} \int_{\partial M} \{\phi n^\mu\nabla_\mu(Af) - (Af)n^\mu\nabla_\mu\phi\}\sqrt{-h} d^3x &= - \int_{\Sigma_{t_1}} \{\phi n^\mu\nabla_\mu(Af) - (Af)n^\mu\nabla_\mu\phi\}\sqrt{-h} d^3x \\ &= \Omega(\Delta f, \phi), \end{aligned} \quad (\text{B.6})$$

where we have used in the last line that Rf vanishes at Σ_{t_1} . \square

Theorem B.3. $G_R(x, z) = G_A(z, x)$.

Proof. Let M be the region bounded by a pair of Cauchy surfaces, one to the future of both x and z , the other to their past. Then

$$\begin{aligned} G_R(x, z) - G_A(z, x) &= \int_M \{G_R(y, z)\delta^4(y - x) - G_A(y, x)\delta^4(y - z)\} d^4y \\ &= \int_M \{G_A(y, x)(\square_y + m^2)G_R(y, z) - G_R(y, z)(\square_y + m^2)G_A(y, x)\} dV_y \\ &= \int_M \{G_A(y, x)\square_y G_R(y, z) - G_R(y, z)\square_y G_A(y, x)\} dV_y \\ &= \int_M \nabla_\mu\{G_A(y, x)\nabla^\mu G_R(y, z) - G_R(y, z)\nabla^\mu G_A(y, x)\} dV_y \\ &= \int_{\partial M} \{G_A(y, x)n^\mu\nabla_\mu G_R(y, z) - G_R(y, z)n^\mu\nabla_\mu G_A(y, x)\}\sqrt{-h} d^3y = 0 \end{aligned}$$

where we have used the Stoke's theorem in the last line. Here n^μ is the unit normal to the boundary ∂M and h is the determinant of the induced metric on the boundary. We now explain why the last expression is identically zero. By definition, it is only when y is in the causal future of z and the causal past of x where this expression could be nonzero. However, because y is being evaluated at the boundary this is never possible in a globally hyperbolic spacetime. \square

References

- [1] S. Johnston, *Feynman propagator for a free scalar field on a causal set*, *Phys. Rev. Lett.* **103** (Oct, 2009) 180401.
- [2] R. D. Sorkin, *Scalar field theory on a causal set in histories form*, *Journal of Physics: Conference Series* **306** (2011), no. 1 012017.
- [3] S. Hawking, *Particle creation by black holes*, *Communications in Mathematical Physics* **43** (1975) 199–220. 10.1007/BF02345020.
- [4] J. J. Bisognano and E. H. Wichmann, *On the duality condition for a hermitian scalar field*, *Journal of Mathematical Physics* **16** (1975), no. 4 985–1007.
- [5] W. G. Unruh, *Notes on black-hole evaporation*, *Phys. Rev. D* **14** (Aug, 1976) 870–892.
- [6] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, *Theory of cosmological perturbations*, *Physics Reports* **215** (1992), no. 5-6 203 – 333.
- [7] B. Allen, *Vacuum states in de sitter space*, *Physical Review D* **32** (1985) 3136–3149.
- [8] R. Wald, *Quantum field theory in curved spacetime and black hole thermodynamics*. Chicago lectures in physics. University of Chicago Press, 1994.
- [9] R. M. Wald, *The Formulation of Quantum Field Theory in Curved Spacetime*, [arXiv:0907.0416](https://arxiv.org/abs/0907.0416).
- [10] F. Dowker, S. Johnston, and R. D. Sorkin, *Hilbert spaces from path integrals*, *Journal of Physics A: Mathematical and Theoretical* **43** (2010), no. 27 275302, [[arXiv:1002.0589](https://arxiv.org/abs/1002.0589)].
- [11] R. D. Sorkin, *Quantum mechanics as quantum measure theory*, *Mod.Phys.Lett.* **A9** (1994) 3119–3128, [[gr-qc/9401003](https://arxiv.org/abs/gr-qc/9401003)].
- [12] R. D. Sorkin, *Toward a 'fundamental theorem of quantal measure theory'*, [arXiv:1104.0997](https://arxiv.org/abs/1104.0997).
- [13] J. B. Hartle, *Space-time quantum mechanics and the quantum mechanics of space-time*, [gr-qc/9304006](https://arxiv.org/abs/gr-qc/9304006). 183pages, UCSB92-21, (corrected macro calls) Journal-ref: in *Gravitation and Quantizations: Proceedings of the 1992 Les Houches Summer School*, ed. by B. Julia and J. Zinn-Justin, North Holland, Amsterdam, 1995.
- [14] A. Ashtekar and A. Magnon-Ashtekar, *A curiosity concerning the role of coherent states in quantum field theory*, *Pramana* **15** (1980) 107–115. 10.1007/BF02847917.
- [15] M. Reed and B. Simon, *I: Functional Analysis, Volume 1 (Methods of Modern Mathematical Physics) (vol 1)*. Academic Press, rev enl su ed., Jan., 1981.
- [16] C. J. Fewster and R. Verch, *On a Recent Construction of 'Vacuum-like' Quantum Field States in Curved Spacetime*, [arXiv:1206.1562](https://arxiv.org/abs/1206.1562).
- [17] N. Afshordi, M. Buck, F. Dowker, D. Rideout, R. D. Sorkin, et al., *A Ground State for the Causal Diamond in 2 Dimensions*, [arXiv:1207.7101](https://arxiv.org/abs/1207.7101).
- [18] N. Afshordi, S. Aslanbeigi, M. Buck, F. Dowker, and R. Sorkin. In preparation.
- [19] S. Fulling, *Remarks on positive frequency and hamiltonians in expanding universes*, *General Relativity and Gravitation* **10** (1979) 807–824. 10.1007/BF00756661.
- [20] C. Dappiaggi, T.-P. Hack, and N. Pinamonti, *Approximate KMS states for scalar and spinor fields in Friedmann-Robertson-Walker spacetimes*, *Annales Henri Poincare* **12** (2011) 1449–1489, [[arXiv:1009.5179](https://arxiv.org/abs/1009.5179)].

- [21] N. Birrell and P. Davies, *Quantum Fields in Curved Space*. Cambridge University Press, Cambridge, 1982.
- [22] F. W. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, NY, USA, 1st ed., 2010.